

ON THE REGISTRABILITY OF TWO CT VOLUMES

Diego Fiorin, Marie-Pierre Jolly, and Charles Florin

Siemens Corporate Research
Imaging and Visualization Department
Princeton, NJ, USA
marie-pierre.jolly@siemens.com

ABSTRACT

This paper describes a method to determine whether two CT volumes overlap in anatomy or not. This is an important problem because radiologists have to manually select which series should be registered together for follow-up exams. This task is becoming more and more tedious as the number of studies and series for each patient increases in large hospital settings, and meta-data is often erroneous, incomplete, or inconsistent, and therefore unreliable. We demonstrate on 40 patients and 100 possible matches that our tool is successful in identifying the overlapping (or registrable) cases automatically. We also show that this is not possible using the residual error after registration.

Index Terms— Computed Tomography, Image Registration

1. INTRODUCTION

Since its introduction in the 1970s, computed tomography (CT) has become an important tool in medical imaging and the gold standard in the diagnosis of a large number of diseases and conditions, including cerebrovascular accidents and intracranial hemorrhage (cranial CT), lung disease, airspace disease, and lung cancer (chest CT), pulmonary embolism and aortic dissection (CT angiography), abdominal disease, cancer staging and follow-up, renal/urinary stones, appendicitis, pancreatitis, bowel obstruction (abdominal and pelvic CT). More information can be found in [1].

Medical image registration is an important field as can be seen from the large amount of work published in the domain [2, 3]. Important categories are rigid vs. non rigid, intensity vs. point based, and intramodal vs. multimodal. Indeed, clinicians are interested in registering datasets acquired with different modalities which highlight different properties of the tissue or organ. They are also interested in registering datasets acquired at different times to follow up on a disease, which is the main focus of this paper.

In the literature, it is always assumed that two volumes being examined by a registration algorithm actually do match.

Diego Fiorin is now at the University of Padova, Italy
Charles Florin is now with Siemens Medical Solutions, Malvern, PA

The goal of the algorithm is then to find the transformation (rigid or not) between the two volumes. Nowadays however, radiologists are more and more confronted with situations where they have large databases with many studies containing many series for the same patient. When a patient comes back for a new exam, it would not make any sense for the current liver series to be registered with a previous lung series.

As a result, the radiologist has to manually select the matching pairs to be sent to the registration algorithm before the follow-up exam can proceed. This procedure can be tedious, especially if the number of studies and series is large. It is important to note that meta-data entered at acquisition time cannot be relied upon. Often, technologists do not spend time entering the correct information, default values are stored, or typos might be inserted accidentally. In addition, the dichotomy of the body might be ambiguous: a scan might be labelled “abdomen” and go from the bottom of the lungs to the pelvis while another scan might be labelled “liver” and go from the bottom of the lungs to the kidneys. These two scans overlap but their meta-data are totally different.

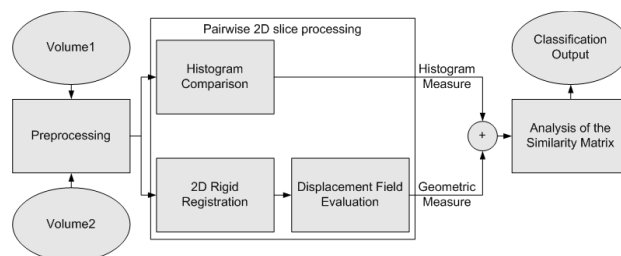


Fig. 1. Flowchart of the method.

The goal of this work is to develop a method to automatically determine whether two volumes should be registered or not. In particular the process should classify the pair as “overlapping” or “registrable” when the two scans cover (even partially) the same anatomy and “non-overlapping” or “not registrable” when the volumes are completely different. It should also give some spatial indication about the matching part between the two volumes, when there is some. To this day, we

are not aware of any work that has addressed this problem. A direct evaluation of the registration residual (i.e.: sum of squared differences, mutual information, ...) cannot be used to make a decision on the registrability of the two volumes as we will see in this paper. Even though the registration procedure aims at optimizing a matching criterion, different runs between different pairs of volumes cannot be compared.

2. PROPOSED METHOD

The approach that we have chosen to determine if two volumes are registrable is two-step. First, slices are compared pairwise to determine how similar they are. This results in a matrix where each entry represents a distance measure between one slice in the first volume and another slice in the other volume. The next step is to analyze this matrix to compute a distance measure between the two volumes. The proposed method is summarized in the flowchart in Figure 1.

2.1. Preprocessing

The algorithm input is composed of two CT volumes. Each scan might have some specific settings (e.g.: slice thickness and pixel size) that depend on the target anatomy and the goal of the exam. For instance, it is common to acquire a routine scan with 5mm slices, while a more precise scan with 1mm slices might be required around a pathology. Thus, the first step in the preprocessing is to resample the volumes so that voxels in both volumes have identical sizes. After this operation, the first volume \mathcal{V}_1 contains N_1 slices $\{\mathcal{S}_1^1, \dots, \mathcal{S}_{N_1}^1\}$ and the second volume \mathcal{V}_2 contains N_2 slices $\{\mathcal{S}_1^2, \dots, \mathcal{S}_{N_2}^2\}$.

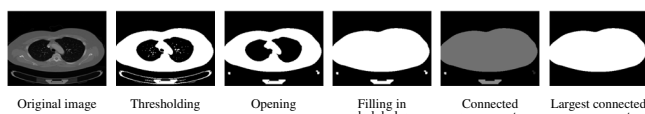


Fig. 2. Preprocessing on one slice to segment the body of interest from the irrelevant background.

The next step is to detect the body outline and remove the table and dark background. For this, we use a combination of very simple image processing steps as illustrated in Figure 2. First, we apply Otsu's thresholding technique [4] to isolate structures from the dark background. Second, we apply an opening morphological operator to eliminate small and thin objects. Third, we fill in dark holes corresponding to air pockets in the lungs or colon. Finally, we retain the largest connected component. This operation produces a mask area \mathcal{M}_i^v for the body in every slice \mathcal{S}_i^v , $i = 1, \dots, N_v$, $v = 1, 2$.

2.2. Histogram comparison

Once the slices have been preprocessed, the goal is to build a $N_1 \times N_2$ distance matrix \mathbf{M} , so that the entry M_{ij} contains the distance cost between slice \mathcal{S}_i^1 from volume \mathcal{V}_1 and slice

\mathcal{S}_j^2 from volume \mathcal{V}_2 . The first distance measure that we chose is based on histogram comparison. There are a large number of histogram distance measures that have been published in the literature (e.g.: intersection, Kullback-Leibler, Jeffrey divergence, χ^2 , Kolmogorov-Smirnov, Bhattacharyya, ...). We eliminated the parametric measures right away since our histograms were far from Gaussian or even unimodal. Among the non parametric measures, we chose to use the L_1 Minkowsky distance because it is not influenced by the size or the absolute gray levels of the regions. Thus, the histogram distance is defined as $M_{ij} = \sum_{k=1}^K |h_i^1(k) - h_j^2(k)|$ where the histogram h_i^v is built over the region of interest \mathcal{M}_i^v of slice \mathcal{S}_i^v , $i = 1, \dots, N_v$, $v = 1, 2$; K is the number of gray levels in the histograms (300 in our case).

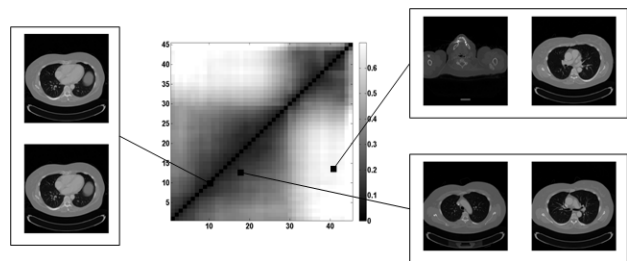


Fig. 3. Distance matrix between two identical volumes based on the Minkowsky histogram distance.

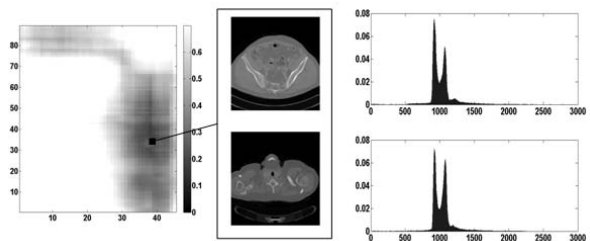


Fig. 4. Low cost region in the similarity matrix corresponding to very different slices having almost identical histograms.

Figure 3 shows the structure of the matrix for two identical volumes. The diagonal that corresponds to identical slices contains the lowest costs. The region near the diagonal is a transition region where the slices are not so different from each other. The cells most distant from the diagonal contain the highest costs corresponding to very different slices.

Unfortunately, the histogram distance is not always discriminant enough. Figure 4 shows an example where two very different slices have very similar histograms and the corresponding region in the distance matrix has low cost. In fact, quite often, the histograms of the shoulder area and the lower intestins/upper pelvis area are very similar. This is due to the fact that histograms do not take into account any spatial information. For this reason, we have introduced another term in the distance matrix based on 2D local registration and the resulting displacement field.

2.3. 2D registration and displacement field

The basic idea behind this component is that if two very different slices are registered non-rigidly, the resulting displacement field should be very large. In order to eliminate the rigid component in the displacement field, all slices are rotated to align the main axis of the mask \mathcal{M} with the horizontal axis and resized to a 256×256 square to remove any scale and translation bias.

To find the displacement field between two slices, we use the algorithm proposed by Chefd'Hotel *et al.* [5] for multimodal image matching. The method uses a variational approach to optimize the cross correlation between the two images combined with a regularization term to control the smoothness of the deformation field. We chose to use the median of the displacement field as the geometric distance measure between the two slices \mathcal{S}_i^1 and \mathcal{S}_j^2 from volumes \mathcal{V}_1 and \mathcal{V}_2 . Indeed, we found that the median was not influenced by outliers as much as the average.

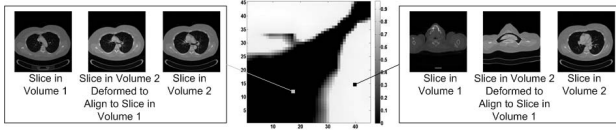


Fig. 5. Distance matrix between two identical volumes based on the displacement field.

Figure 5 shows the structure of the matrix for the same two identical volumes as in Figure 3. For two particular points in the matrix, it shows the original slices as well as the slice from \mathcal{V}_2 that was deformed to align to the slice from \mathcal{V}_1 . The two slices on the left are fairly similar, so both the registration residual and the displacement field are small. The two slices on the right are very different, but the registration residual is small because the registration algorithm is given enough freedom to deform one slice into the other. The displacement field however is very large, and that is the reason why we decide to use it as our distance measure.

2.4. Alignment using the distance matrix

The final distance matrix is generated by adding the histogram distances to the median values of the displacement field (normalized between 0 and 1). Let us now examine the structure of the distance matrix. When two volumes contain some matching part, a low cost canal surrounded by some higher cost area appears in the distance matrix. Some of these situations are shown in Figure 6. Since the correspondence between the two volumes is more or less a translation in the z-direction, we model it as a straight line through the distance matrix. We define the cost of a straight line as the sum of the entries in the matrix along the line, averaged over the length of the line. We perform an exhaustive search to find the straight line with minimum cost which corresponds to the best alignment between the two volumes. In addition, we impose some

constraints on the position and orientation of the line. For example, the slope of the line has to be positive since both volumes are oriented from head to feet. Also, enough slices should be considered, so the slope of the line cannot be too close to 0 or to ∞ . In a nutshell, the starting points for the line can be anywhere in the first $2/3^{\text{rd}}$ of either volume, while the end points can be anywhere in the last $2/3^{\text{rd}}$.

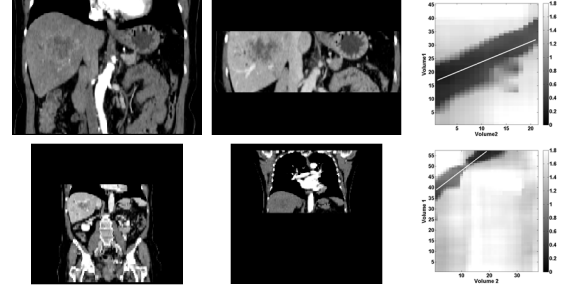


Fig. 6. Relationship between the structure of the distance matrix and the corresponding volumes (the best line is also shown on top of the matrix).

3. EXPERIMENTS

Our test dataset includes 40 different patients each containing two studies taken at different times. 20 of the patients have one series in each study (2 volumes per patient) and another 20 more patients have two series in each study (4 volumes per patient). Since we are only interested in matching datasets from the same patient, this yields 100 possible matches. Among these 100 matches, 51 are known to be registrable and the remaining 49 are not registrable. Note that many of these datasets came from oncology patients and therefore contain tumors. Their presence however did not seem to affect the performance of our algorithm.

We applied the algorithm to all 100 possible matches and computed the distance between the two volumes (cost of the best straight line through the distance matrix). Figure 7(a) shows the distributions of these distances for the 51 registrable examples and the 49 non-registrable examples. The bottom of the figure shows the box and whiskers plot representation of the samples while the top of the figure shows a Gaussian modelling of the two distributions. Information about the mean and standard deviation of the distributions are shown in Table 1(a). We chose a threshold at the intersection of the two Gaussian distributions to classify the samples and obtained the classification errors P_{FP} of false positive (type I error), P_{FN} of false negative (type II error), and P_{ERR} of total error as reported in Table 1(a). It can be seen that the distribution are very well separated and the classification errors are very low.

For comparison, we also registered the 100 volume pairs using normalized mutual information (NMI) as described in [6] where it was established that NMI was the best entropy

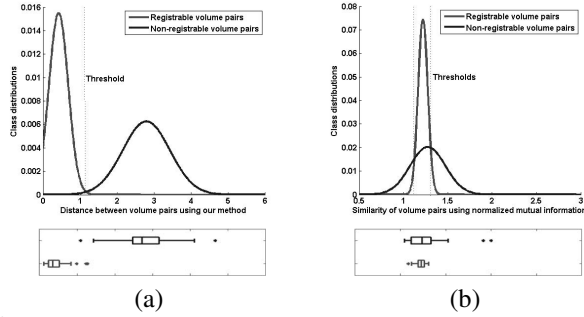


Fig. 7. Distributions of the 51 registrable and 49 non-registrable volume pairs represented by a box and whiskers plot and a Gaussian model for (a) the distance between two volumes using our method and (b) the similarity between two volumes using NMI registration.

(a)		μ	σ	$T = 1.16$	P_{FP}	0.53%
	ω_0	0.42	0.26			
	ω_1	2.79	0.64	$x < T \Rightarrow \omega_0$	P_{FN}	0.17%
				$x > T \Rightarrow \omega_1$	P_{ERR}	0.70%
(b)		μ	σ	$T_1 = 1.13$ $T_2 = 1.31$	P_{FP}	10.7%
	ω_0	1.22	0.05			
	ω_1	1.27	0.20	$T_1 < x < T_2 \Rightarrow \omega_0$	P_{FN}	3.4%
				$x < T_1 \Rightarrow \omega_1$	P_{ERR}	24.1%
				$x > T_2 \Rightarrow \omega_1$		

Table 1. Statistics on the distributions ω_0 of overlapping pairs and ω_1 of non-overlapping pairs, decision rules, and probability of false positives P_{FP} , false negatives P_{FN} , and error P_{ERR} for (a) our method and (b) NMI registration.

based measure to align 3D medical volumes. The distributions of the NMI values after registration for the 51 registrable and 49 non-registrable pairs are shown in Figure 7(b). The parameters of the Gaussian distributions, thresholds, decision rules, and classification errors are shown in Table 1(b). It can be seen that the two distributions completely overlap, the classification errors are very large, and the classes cannot be separated at all using NMI registration. This clearly demonstrates the superiority of our method.

To evaluate the performance of our algorithm at generalizing, we used k -fold cross validation which consists of dividing the samples into k groups where the samples in $k - 1$ groups are used for training while the samples in the remaining group are used for testing. In our case, the folds were selected around the patients so that no patient appeared in both the training and the test set. We decided to use 10 folds, each containing 4 patients. The error rate was estimated using Monte Carlo sampling with 500 iterations [7]. In our problem, we want to avoid false negatives, because that would mean missing true registrable cases and having the radiologist spend time looking for them in the database. So we did not choose the threshold as the intersection of the two Gaussians. Instead, we chose the threshold so as to minimize the false negative rate. The results of the k -fold cross validation experiments are shown in Table 2. It can be seen that the average error rate is just above 2% with a very low false negative rate.

	0	1	2	3	4	average
P_{FP}	-	0.6%	80.8%	18.6%	-	2.180
P_{FN}	99.4%	0.6%	-	-	-	0.006
P_{ERR}	-	0.6%	80.4%	18.8%	0.2%	2.186

Table 2. Probability of false positives P_{FP} , false negatives P_{FN} , and error P_{ERR} for 500 k -fold cross validation.

4. CONCLUSION

Patient databases are getting larger and larger and the task of exam follow-up is getting more and more tedious for the radiologist who has to manually determine which series cover the same part of the anatomy and should be registered together. In many cases, the meta-data associated with the images cannot be used to make this decision automatically because it is erroneous, incomplete, or inconsistent. In this paper, we tackle the new problem of trying to determine automatically whether two CT scans overlap and should be registered. We have demonstrated on a large number of patients (40) and possible matches (100) that we were able to differentiate registrable and non-registrable datasets. We have also shown that this kind of decision cannot be made using a classical registration algorithm (such as NMI) because, even though the process optimizes some criterion, outputs from different runs cannot be compared.

In the future, we would like to use the output of our algorithm, namely the straight line, to initialize the registration process. This would most likely guarantee a better, faster, and more accurate convergence of the registration optimization process. Also, the process needs to be tested on a very large set of examples to fully evaluate the performance of the system, especially for longitudinal studies with pathologies that vary a lot with time. Finally, we would like to extend our system to consider multi-modality problems.

5. REFERENCES

- [1] American College of Radiology and Radiological Society of North America, "RadiologyInfo," www.radiologyinfo.org.
- [2] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998.
- [3] J. M. Fitzpatrick, D. L. G. Hill, and Jr. C. R. Maurer, "Image registration," in *Handbook of Medical Imaging. Volume 2. Medical Image Processing and Analysis*, pp. 447–514. SPIE Press, 2000.
- [4] N. Otsu, "A threshold selection method from gray level histograms," *IEEE Trans. Systems, Man and Cybernetics*, vol. 9, pp. 62–66, 1979.
- [5] C. Chef'd'Hotel, G. Hermosillo, and O. Faugeras, "Flows of diffeomorphisms for multimodal image registration," in *Proc. IEEE Int. Symp. Biomedical Imaging*, 2002, pp. 753–756.
- [6] C. Studholme, C. L. G. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognition*, vol. 32, no. 1, pp. 71–86, 1999.
- [7] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proc. Int. Joint Conference on Artificial Intelligence*, 1995, pp. 1137–1145.